

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 09-091257

(43)Date of publication of application : 04.04.1997

(51)Int.Cl.

G06F 15/16

G06F 9/46

(21)Application number : 07-270513

(71)Applicant : NEC CORP

(22)Date of filing : 25.09.1995

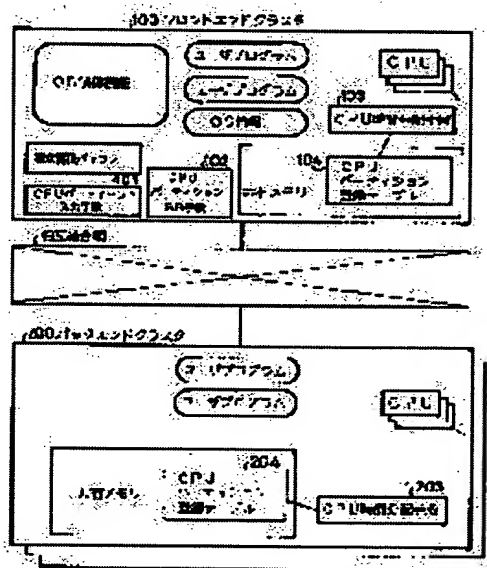
(72)Inventor : TANAKA HIDETOSHI

(54) CPU MANAGEMENT SYSTEM

(57)Abstract:

PROBLEM TO BE SOLVED: To optimize the use of CPUs by dividing CPUs into groups to manage them as CPU partitions in a multi-cluster system.

SOLUTION: A CPU partition input means 101 inputs CPU partition information related to respective clusters (a front end cluster 100 and back end clusters 200). A CPU partition SG means 102 divides CPU groups of respective clusters into groups based on CPU partition information and registers information indicating group division in CPU partition registration tables 104 and 204 in each cluster. At the time of the occurrence of a process, a CPU time distribution means 103 or 203 assigns the CPU time related to CPUs belonging to the CPU partition, which is adapted to the character of the process, to the process based on reference to the CPU partition registration table 104 or 204.



LEGAL STATUS

[Date of request for examination] 25.09.1995

[Date of sending the examiner's decision of rejection] 14.10.1998

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

BEST AVAILABLE COPY

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-91257

(43) 公開日 平成9年(1997)4月4日

(51) Int.Cl. ⁸	識別記号	庁内整理番号	FI	技術表示箇所
G 0 6 F 15/16	3 8 0		G 0 6 F 15/16	3 8 0 Z
9/46	3 6 0		9/46	3 6 0 B

審査請求 有 請求項の数 4 FD (全 10 頁)

(21) 出願番号 特願平7-270513

(22) 出願日 平成7年(1995)9月25日

(71) 出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72) 発明者 田中 英俊

東京都港区芝五丁目7番1号 日本電気株式会社内

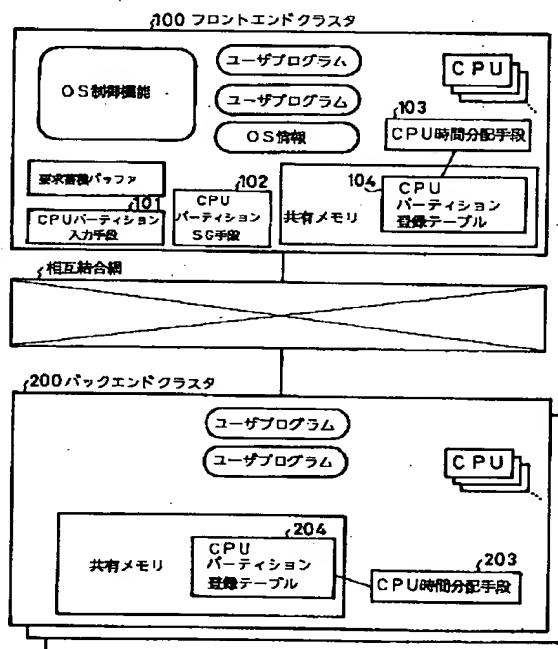
(74) 代理人 弁理士 河原 純一

(54) 【発明の名称】 CPU管理方式

(57) 【要約】

【課題】 マルチクラスタシステムにおいて、CPUをグループ化してCPUパーティションとして管理し、CPUの利用の最適化を図る。

【解決手段】 CPUパーティション入力手段101は、各クラスタ(フロントエンドクラスタ100およびバックエンドクラスタ200の各々)に関するCPUパーティション情報を入力する。CPUパーティションSG手段102は、CPUパーティション情報に基づき各クラスタのCPU群のグループ分けを行い、そのグループ分けを示す情報を各クラスタ内のCPUパーティション登録テーブル104および204に登録する。CPU時間分配手段103または203は、プロセスの発生時に、CPUパーティション登録テーブル104または204の参照に基づき、当該プロセスの性格に適合するCPUパーティションに属するCPUに係るCPU時間を当該プロセスに割り当てる。



【特許請求の範囲】

【請求項1】 1つのフロントエンドクラスタを含む複数のクラスタを相互結合網で接続したマルチクラスタシステムにおけるCPU管理方式において、

各クラスタに関するCPUパーティション情報を入力するフロントエンドクラスタ内のCPUパーティション入力手段と、

前記CPUパーティション入力手段により入力されたCPUパーティション情報に基づいて各クラスタにおける複数のCPUのグループ分けを行い、そのグループ分けを示す情報を各クラスタ内のCPUパーティション登録テーブルに登録するフロントエンドクラスタ内のCPUパーティションSG手段と、

プロセスの発生時に前記CPUパーティション登録テーブルの参照に基づいて当該プロセスの性格に適合するCPUパーティションに属するCPUに係るCPU時間を当該プロセスに割り当てる各クラスタ内のCPU時間分配手段とを有することを特徴とするCPU管理方式。

【請求項2】 自クラスタにおける各CPUパーティションの使用率を自クラスタ内のCPUパーティション使用率情報テーブルに登録する各クラスタ内のCPU時間分配手段と、

前記CPUパーティション使用率情報テーブルに基づいて使用率の低いCPUパーティションに属する空きCPUを使用率の高い別のCPUパーティションに動的に割り当て、各CPUパーティションの使用率の最適化を図るCPUパーティションSG手段とを有することを特徴とする請求項1記載のCPU管理方式。

【請求項3】 CPUパーティション使用率情報テーブルが「パーティション番号」と「使用時間」との対応情報を有することを特徴とする請求項2記載のCPU管理方式。

【請求項4】 自クラスタ内のCPUパーティション使用率情報テーブルに基づいて使用率が適切でないCPUパーティションに関する警告を出力する各クラスタ内のCPU時間分配手段を有することを特徴とする請求項2または3記載のCPU管理方式。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、マルチクラスタシステムにおけるCPU(Central Processing Unit)管理方式に関する。

【0002】ここで、本願明細書でいう「マルチクラスタシステム」とは、以下の①～④に示すような特徴を有するシステムをいう。

① 複数のクラスタ(複数のCPUとそれらの共有メモリとから構成される独立した単位)によって構成される。

② OS(Operating System)制御機能、OS情報、およびユーザプログラムを有し(これら

の情報は共有メモリに格納されている)、バックエンドクラスタからの要求を蓄積する要求蓄積バッファを持ち、マルチクラスタ環境の中に1つだけ存在可能なフロントエンドクラスタを有する。

③ OS制御機能を持たずに、ユーザプログラムを有し(ユーザプログラムの情報は共有メモリに格納されている)、マルチクラスタ環境に複数個存在可能なバックエンドクラスタを有する。

④ 複数のクラスタ(1つのフロントエンドクラスタおよび1以上のバックエンドクラスタ)が相互結合網(一般的には高速ネットワーク)で接続されている。

【0003】

【従来の技術】従来より、複数のCPUを有するコンピュータシステムにおいては、CPUをグループ化して、グループに関する情報に基づいてCPUの管理を行うCPU管理方式が存在した(例えば、特開平1-109464号公報参照)。

【0004】しかし、従来のCPU管理方式では、上述のようなマルチクラスタシステムに関しては、対処することができなかった(上述の特開平1-109464号公報においてもマルチクラスタシステムについては言及されていない)。すなわち、マルチクラスタシステムにおける従来のCPU管理方式では、グループ化の考え方は導入されていなかった。

【0005】

【発明が解決しようとする課題】上述したように、従来のCPU管理方式では、マルチクラスタシステムに適切に対処することができず、CPUをグループ化することによる効果(各種性能の向上)をマルチクラスタシステムにおいて実現することができないという問題点があった。

【0006】本発明の目的は、上述の点に鑑み、マルチクラスタシステムにおいて、使用目的(ジョブ内容)毎にCPUをグルーピング(グループ化)してCPUパーティションとして管理することにより、CPUの利用の最適化を図り、当該マルチクラスタシステムにおいてCPUのグループ化による各種性能(マルチクラスタシステムにおけるTSS(Time Sharing System)の応答性、多重バッチジョブのスループット、および並列プログラムに関するジョブのTAT(Turn Around Time)等の性能)の向上を達成することができるCPU管理方式を提供することにある。

【0007】

【課題を解決するための手段】本発明のCPU管理方式は、1つのフロントエンドクラスタを含む複数のクラスタを相互結合網で接続したマルチクラスタシステムにおけるCPU管理方式において、各クラスタに関するCPUパーティション情報を入力するフロントエンドクラスタ内のCPUパーティション入力手段と、前記CPUパ

10

20

30

40

50

3

ーティション入力手段により入力されたCPUパーティション情報に基づいて各クラスタにおける複数のCPUのグループ分けを行い、そのグループ分けを示す情報を各クラスタ内のCPUパーティション登録テーブルに登録するフロントエンドクラスタ内のCPUパーティションSG手段と、プロセスの発生時に前記CPUパーティション登録テーブルの参照に基づいて当該プロセスの性格に適合するCPUパーティションに属するCPUに係るCPU時間を当該プロセスに割り当てる各クラスタ内のCPU時間分配手段とを有する。

【0008】

【発明の実施の形態】次に、本発明について図面を参照して詳細に説明する。

【0009】第1に、本発明のCPU管理方式の第1の実施例(請求項1記載の発明に対応する実施例)について説明する。

【0010】図1は、本発明のCPU管理方式の第1の実施例の構成を示すブロック図である。

【0011】本実施例のCPU管理方式は、1つのフロントエンドクラスタ100と複数のバックエンドクラスタ200とが相互結合網(高速ネットワーク)で接続されているマルチクラスタシステムにおいて実現される。

【0012】フロントエンドクラスタ100は、OS制御機能、OS情報、およびユーザプログラム(図示していないが、これらの情報は共有メモリに格納されている)や、要求蓄積バッファ、複数のCPU、および共有メモリを備えた上で、CPUパーティション入力手段101と、CPUパーティションSG(System Generation)手段102と、CPU時間分配手段103と、CPUパーティション登録テーブル104(共有メモリ内に存在する)とを含んで構成されている。

【0013】バックエンドクラスタ200は、複数のCPU、ユーザプログラム、および共有メモリを備えた上で、CPU時間分配手段203と、CPUパーティション登録テーブル204(共有メモリ内に存在する)とを含んで構成されている。

【0014】本実施例のCPU管理方式は、図1に示すようなマルチクラスタシステム内の各クラスタにおいて、複数のCPUを使用目的(ジョブ内容)毎にグルーピングしてCPUパーティションとして管理する。すなわち、各クラスタ(フロントエンドクラスタ100およびバックエンドクラスタ200の各々)における複数のCPUをグルーピングしてそのグループ分けを示す情報を各クラスタに対応するCPUパーティション登録テーブル104または204に登録することにより、CPUパーティションを特定の使用目的(ジョブ内容)毎に管理できるようにする。

【0015】図2は、本実施例のCPU管理方式の具体的な動作を説明するための図である。

4

【0016】図3は、本実施例のCPU管理方式の処理を示す流れ図である。この処理は、CPUパーティション情報入力ステップ301と、CPUパーティション登録テーブル作成ステップ302と、プロセス発生判定ステップ303と、CPU選択ステップ304と、CPU時間割当てステップ305とからなる。

【0017】次に、このように構成された本実施例のCPU管理方式の動作について説明する。

【0018】フロントエンドクラスタ100内のCPUパーティション入力手段101は、システム管理者からの指示に基づき、CPUパーティションの性格付けを行いその管理を行うための情報であるCPUパーティション情報を、各クラスタ(フロントエンドクラスタ100およびバックエンドクラスタ200の各々)について入力する(ステップ301)。ここで、本実施例におけるCPUパーティション情報とは、具体的には、ジョブ内容(CPUバウンド(bound)ジョブ、インタラクティブ(interactive)ジョブ、およびバッチジョブ等)によって分類されたCPUパーティションと、そのCPUパーティションに属するCPUの個数との、対応情報をいう。

【0019】CPUパーティションSG手段102は、ステップ301で入力されたCPUパーティション情報に基づき、各クラスタにおけるCPUパーティション登録テーブル104および204を作成する(ステップ302)。

【0020】図2の例におけるCPUパーティション登録テーブル104では、CPU「3」、「5」、および「12」(CPUの識別情報が「3」、「5」、および「12」であるCPU)が1つのグループとされ、それらの識別情報がCPUパーティション「1」(CPUバウンドパーティション)に対応する情報として格納されている。また、CPU「17」、「18」、「19」、および「20」が1つのグループとされ、それらの識別情報がCPUパーティション「2」(インタラクティブパーティション)に対応する情報として格納されている。さらに、CPU「7」、「9」、「22」、および「23」が1つのグループとされ、それらの識別情報がCPUパーティション「3」(バッチジョブパーティション)に対応する情報として格納されている。このような情報の格納(登録)が、CPUパーティション登録テーブル104ばかりでなく、各バックエンドクラスタ200内のCPUパーティション登録テーブル204についても行われる。これにより、静的に、CPUの使用目的(ジョブ内容)毎にCPUの識別情報がグループ化されることになる。

【0021】フロントエンドクラスタ100内のCPU時間分配手段103または各バックエンドクラスタ200内のCPU時間分配手段203は、当該マルチクラスタシステムの実行時において、プロセスが生じると(ス

5

テップ303参照)、そのプロセスへのCPU時間の割当てを、以下の①および②に示すようにして行う。

【0022】① CPUパーティション登録テーブル104内のCPUパーティションに関する情報に基づいて、割当て対象のプロセスの性格(使用目的、すなわち当該プロセスに関するジョブ内容)に適合したCPUパーティションに属する空きCPUの1つを選択する(ステップ304)。

【0023】② ①で選択したCPUを一定時間そのプロセスに割り当て、すなわちそのCPUに係るCPU時間
10をそのプロセスに割り当て(ステップ305)、その後ステップ303の判定を行う状態に戻る。

【0024】以上のようにして、当該マルチクラスタシステム内の多数のCPUをグルーピングしてCPUパーティションとして管理することができるようになる。

【0025】このようなCPUパーティション(CPUのグループ化)の考え方を導入することにより、マルチクラスタシステムにおける各種性能、すなわちTSSの応答性、多重バッチジョブのスループット、および並列プログラムに関するジョブのTAT等の性能に関して、
20CPU時間の分配の最適化を図ることができる。

【0026】第2に、本発明のCPU管理方式の第2の実施例(請求項2~4記載の発明に対応する実施例)について説明する。

【0027】図4は、本発明のCPU管理方式の第2の実施例の構成を示すブロック図である。

【0028】本実施例のCPU管理方式は、第1の実施例に比べて、フロントエンドクラスタ100内のCPUパーティション使用率情報テーブル105および各バックエンドクラスタ200内のCPUパーティション使用
30率情報テーブル205が追加されている。

【0029】図5は、本実施例のCPU管理方式の具体的な動作を説明するための図である。

【0030】次に、このように構成された本実施例のCPU管理方式の動作について説明する。なお、CPUパーティション使用率情報テーブル105および205に関する動作以外の動作については第1の実施例における動作と同様であるので、以下では言及しない。

【0031】本実施例のCPU管理方式では、CPUパーティションの使用率(当該CPUパーティションに属
40するCPUの使用率)がCPUパーティション使用率情報テーブル105および205に登録される。本実施例では、各CPUパーティション(パーティション番号で識別されるCPUパーティション)に属する全CPUの使用時間によって「CPUパーティションの使用率」が表現されるものとする。すなわち、図5に示すように、CPUパーティション使用率情報テーブル105(CPUパーティション使用率情報テーブル205も同様)は、「パーティション番号」と「使用時間」との対応情報を有している。なお、CPUパーティション使用率情
50

6

報テーブル105および205内の情報は、CPUパーティションの使用率を示す情報であれば、図5に示す情報に限られるものではない。

【0032】第1の実施例に関する説明で述べたように、各クラスタにおけるCPU時間分配手段103および203は、当該マルチクラスタシステムの実行の過程で、CPUパーティション登録テーブル104および204を参照した上で、発生したプロセスにCPU時間を振り分ける。例えば、バッチジョブの処理を行うプロセスにはCPUパーティション「3」に属するCPUを一定時間割り当てる。

【0033】本実施例では、さらに、CPU時間分配手段103および203は、上述のようなプロセスへのCPU時間の割当て動作の直後に、CPUパーティション使用率情報テーブル105および205内の該当するエントリ(割り当てられたCPUが属するCPUパーティションのパーティション番号を有するエントリ)中の「使用時間」の項目に割り当てたCPU時間の値を加算する。上述の例では、パーティション番号「3」のエントリ中の「使用時間」における時間値に割り当てたCPU時間の値を加算する。

【0034】CPUパーティションSG手段102は、CPUパーティション使用率情報テーブル105および205を利用して、各クラスタにおけるCPUパーティションへのCPUの割当て(CPUのグループ分け)を動的に変更(最適化)する。

【0035】ここでは、当該マルチクラスタシステムにおいて、フロントエンドクラスタ100における多重バッチジョブのスループットの向上を図ることを目的として、フロントエンドクラスタ100におけるCPUパーティション「3」に属するCPUの個数を動的に増加させる場合について考える。

【0036】この場合には、図2におけるフロントエンドクラスタ100におけるCPUパーティション「3」に属するCPUの個数を増やすために、その増加の旨を示す指示がCPUパーティション入力手段101を介してシステム管理者から行われる。なお、この指示は、例えば、CPUパーティション使用率情報テーブル105に基づいてCPU時間分配手段103が出力する警告(CPUパーティション「3」の使用時間が過大であることを示す警告)に基づいて、システム管理者によって行われる(CPU時間分配手段103およびCPU時間分配手段203の各々は、自クラスタ(自己が存在するクラスタ)内のCPUパーティション使用率情報テーブル105または205に基づいて、使用率が適切でないCPUパーティションに関する警告を出力することができる)。

【0037】CPUパーティションSG手段102は、その指示に基づいて、CPUパーティション使用率情報テーブル105によって示されるCPU使用状況を参照

7

して、使用率が低いCPUパーティション「2」から空きCPUであるCPU「20」をCPUパーティション「3」に移動し、CPUパーティション登録テーブル104を図2の状態から図5の状態に変更する。

【0038】これにより、CPUパーティションの使用効率の最適化を図ることができ、この場合であると、多重バッチジョブのスループットの向上を実現することができる。

【0039】

【発明の効果】以上説明したように本発明は、マルチクラスタシステムにおいて、CPUパーティション入力手段、CPUパーティションSG手段、CPU時間分配手段、およびCPUパーティション登録テーブルを設け、各クラスタにおいてCPUをグルーピングしてCPUパーティションとして管理することにより、プロセスの性格（使用目的、すなわちジョブ内容）に適合するようにCPU時間の割当てを行うことができ、マルチクラスタシステムにおける各種性能、すなわちTSSの応答性、多重バッチジョブのスループット、および並列プログラムに関するジョブのTAT等の性能の向上を達成することができるという効果を有する。

【0040】また、上記に加えて、CPUパーティション使用率情報テーブルを設けることにより、CPUパー

8

ティションの使用率とプロセスへのCPU時間の割当てとを調整することができ、各種性能の向上を追及しつつ、CPUの使用効率の向上をさらに図ることができるという効果がある。

【図面の簡単な説明】

【図1】本発明のCPU管理方式の第1の実施例の構成を示すブロック図である。

【図2】図1に示すCPU管理方式の具体的な動作を説明するための図である。

【図3】図1に示すCPU管理方式の処理を示す流れ図である。

【図4】本発明のCPU管理方式の第2の実施例の構成を示すブロック図である。

【図5】図4に示すCPU管理方式の具体的な動作を説明するための図である。

【符号の説明】

101 CPUパーティション入力手段

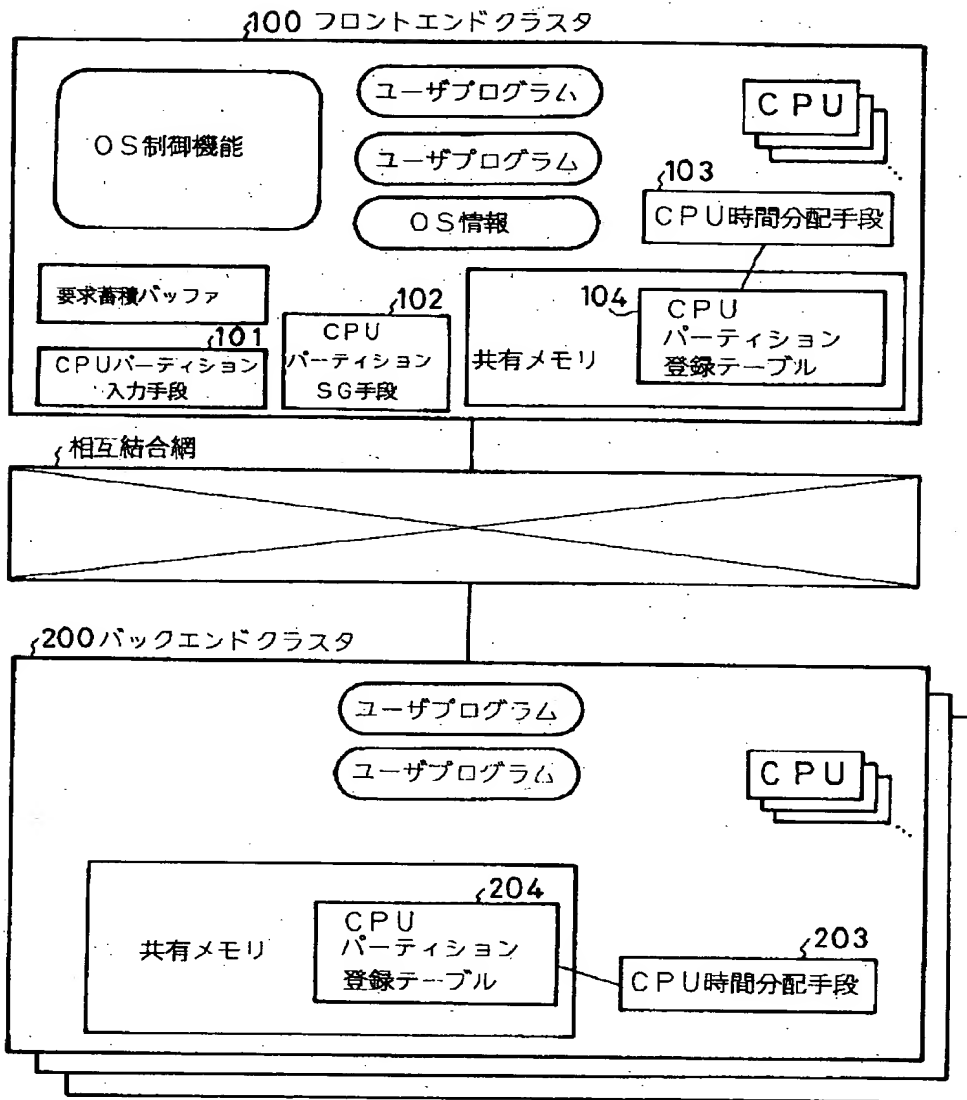
102 CPUパーティションSG手段

103, 203 CPU時間分配手段

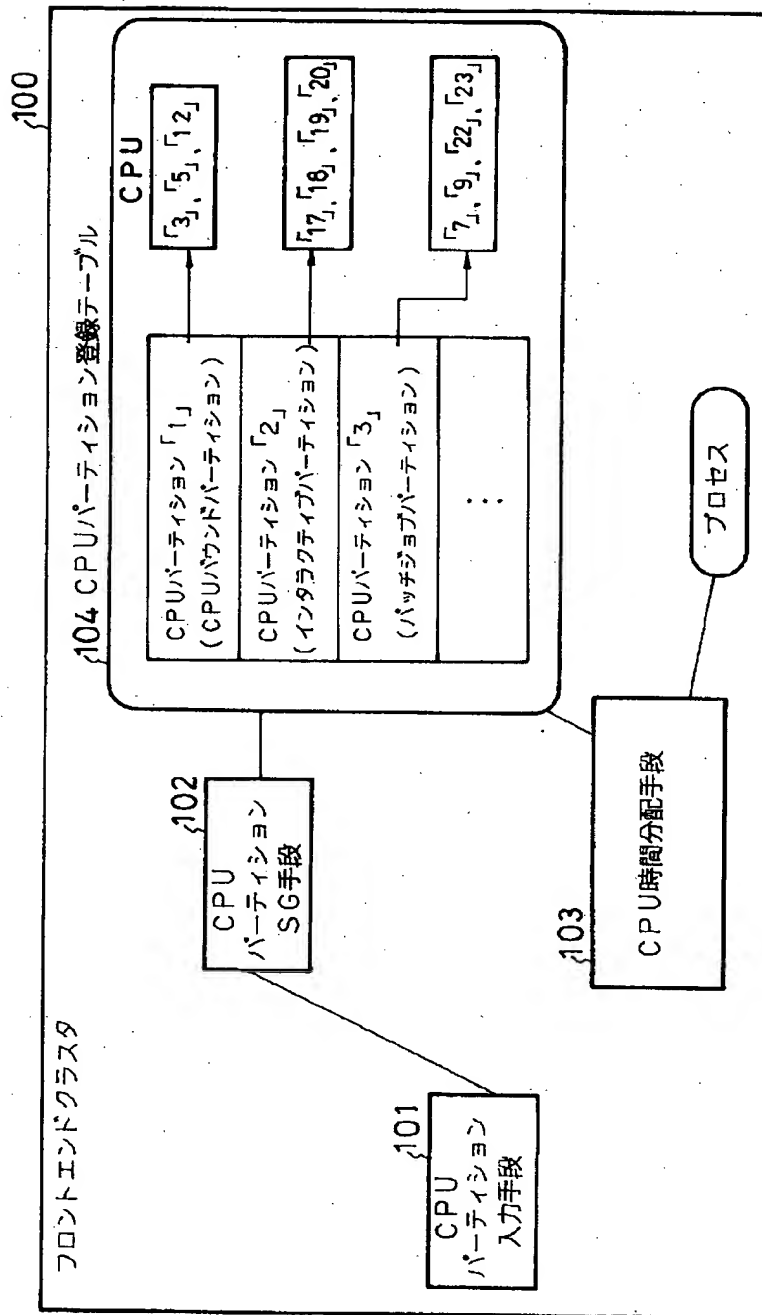
104, 204 CPUパーティション登録テーブル

105, 205 CPUパーティション使用率情報テーブル

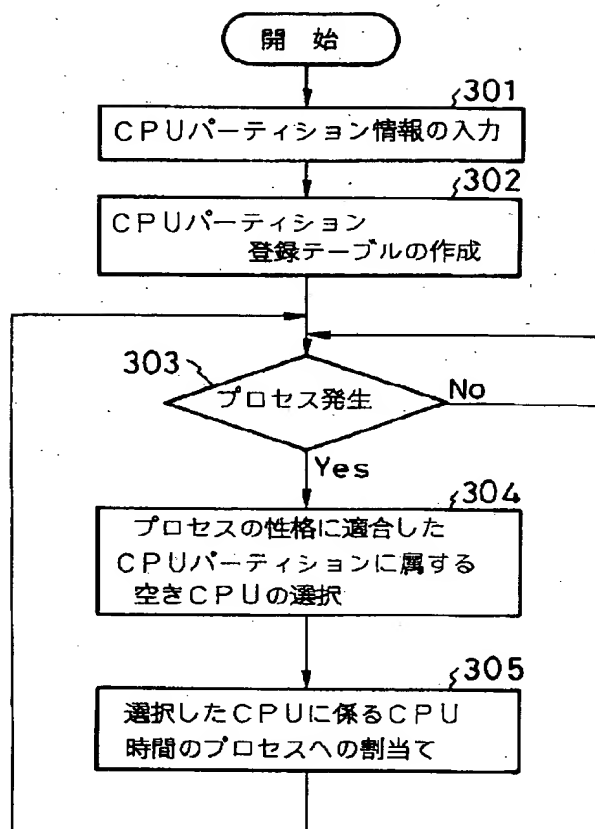
【 図1 】



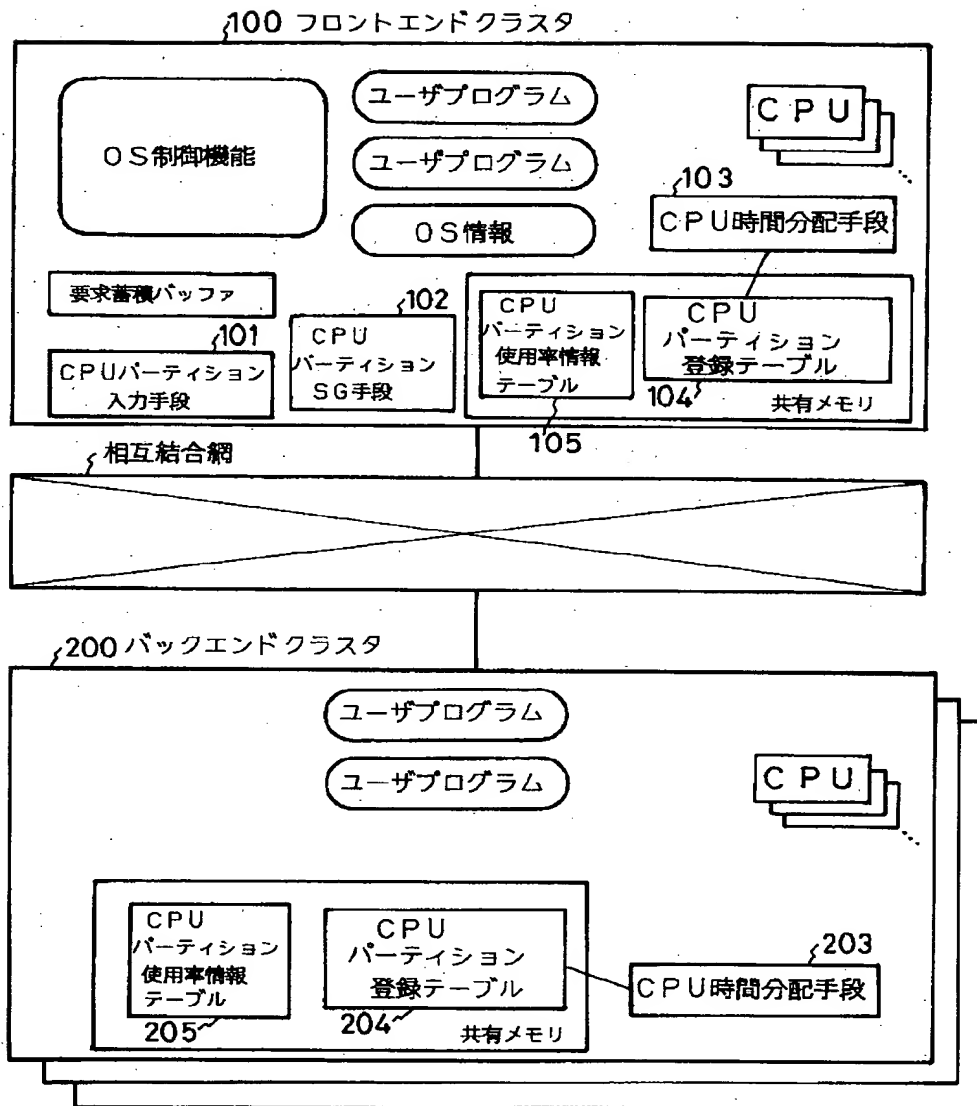
【 図2 】



【 図3 】



【 図4 】



【 図5 】

